Submitted on: 5/30/2014

Yes, we can , if we take over future tasks!

Andreas Degkwitz
Library, Humboldt University, Berlin, Germany.
andreas.degkwitz@ub.hu-berlin.de

Abstract:

Knowledge management is closely connected with the publications of research results, which are the resources for the circulation of knowledge further on. But the linear, item oriented features of printed materials or electronic emulations of printed publications like PDFs are completely different from the rather liquid , digital objects and the networked flows we are observing now. In research environments, which exploit the potential of digital media more consequently than researchers did before, the tasks of knowledge management are faced with completely new challenges. Scholarly communication is changing fundamentally and quite rapidly by the new media and by the tools, which are intensively used in the web-based collaboration and working environments of many disciplines. What causes or constitutes these changes and what are the new roles of the libraries as an essential part of the academic support? How do libraries and information providers meet the new requirements of knowledge management? New approaches of knowledge management will be outlined considering the recent results of the project Future Publications in the Humanities (Fu-PusH), which is granted by the German Research Society (Deutsche Forschungsgemeinschaft) http://www.ub.hu-berlin.de/fu-push-en

Keywords: enhanced publications, information infrastructure, knowledge management, library tasks, value chain.

1. Lost in Traditions

From our everyday experiences the impact of libraries is decreasing significantly. What we are facing, is a complete change of the information behavior since people have their own I-phones or I-pads. With these or other devices you have access to the knowledge of the world in your vest pocket. People are permanently online, searching any kind of contents, reading books or newspapers, looking movies and pictures, joining social networks, writing in blogs, chats, social media and wikis. We are consuming and sharing knowledge and any kind of materials and we are so called multiple pro-sumer , who use, read, annotate, complete, generate and share articles, books, messages, music, movies and whatever you can imagine. To sum it up: All the potential, that digital facilities might have, is being exploited by billons

of users every day. But where are the libraries in our days claiming to be the gateways for information and for the memory of all the universal knowledge? Do libraries play any role in our search engine driven decade, in the so called information or knowledge society? What about the domains of knowledge management librarians are designated and skilled since many centuries? Are they vanished or are they losing their entire awareness?

Many of us might think perhaps or are even convinced, that the impact of the digital age is completely different in the academia, where the diligent and serious approaches of scientific work are still appreciated and existing further on without all the fancy overload of the digital knowledge chunks. From a librarians view there might be assumed or even hoped, that the current patterns of research communication won t go beyond PDF that is nothing else than a more or less useful emulation of printed texts. Therefore with PDF the tradition of all the familiar features of scholarly communication will not change at all and will be carried on for longer. But how are researchers and teachers working actually? Is there a significant gap between the digital day to day life and the approaches to digital means and tools for education and research? We watch the increase of technology and data driven research in many disciplines. Scholars are using all the available mobile devices for communication as well as for producing and sharing knowledge. Virtual research environments and networked publishing is getting more and more relevant. Courseware and learning platforms are part of the most curricula and are established as educational resources. We can observe the ongoing development of so many blogs and wikis and social media, which seem to overtake the tasks of knowledge management the academic libraries are claiming for.

Mostly the current publications and their life cycles, which libraries are managing and providing from their long tradition, is what s left in the basket finally. In other words: All the traditional activities around acquisition, cataloguing and long term access of digital publications are also the key tasks of knowledge management in future? Off course these activities are often completed and enhanced by additional services which are adopted pretty well in many libraries like support for open access publishing, retro digitization of items of the cultural heritage, information literacy, licensing e-books and e-journals that is good and the least! But do all these activities really change the traditional paradigm of librarianship towards an academic support that is appropriate to the upcoming features of data driven research and to the digital day to day life in science and research worldwide? Are the traditional tasks around publications the main issue of the libraries knowledge management in future?

2. Changing Challenges

What s going on in our universities? On the one hand we have the many fields of natural science like biology, chemistry, engineering, informatics, life and earth sciences, medicine, physics etc. These disciplines have in common to deal with data by their empirical and experimental approaches, that information technology is deployed there since many years in order to produce, to process, to evaluate and to aggregate the findings of theses disciplines. Considering the number of existing authoring and publishing tools, which are available, researchers of the mentioned fields don t have any problems to upload the publications of their research results as PDF on an open access server. The missing links might be the metadata of the publications and the management of references, which libraries are able to support, if it doesn t work by default. In case of journal publications the authors will deliver their pre-prints to the publisher, who is recommended in regard of journal ranking and impact. If the articles or even the monographs are licensed, the publications are available based on the contracts with the libraries. However these tasks of libraries won t improve their awareness significantly. But now new kinds and patterns of publications are coming up,

which will change the entire publishing process with new procedure and new tasks of the infrastructure. Before we go more in detail about it we turn the view to the developments in the humanities and the social sciences.

In the empirical and in the experimental sciences the data which are produced and evaluated during the research project have to be aggregated and visualized by software tools. Otherwise the data can t be read or understood software turns the generated data to readable objects. What does happen in the so called Digital Humanities , which offer new and technology driven methods to analyze and to enrich text oriented materials, but also pictures and photographies? In opposite to data non-structured text is neither interoperable nor processible. But if you want to process and evaluate textual materials with digital tools you have to prepare or to transform the material to machine readable data and that means to - e. g. XML or RDF - structured texts. With audio files, digitized material, movies, pictures and photographies you can proceed in similar ways. In other words: Picture and text based material has to be transformed in machine readable data, which can t or can hardly be read by human readers anymore.

Which opportunities are given or offered by the approaches of the Digital Humanities ? Basically we can observe the facilities of annotating, enriching, linking, mining, processing and searching textual or pictural materials, which are either produced for this purpose or are already available for any usage like collections of digitized books or pictures. Here are some examples given explicitly for disciplines of the humanities:

- To mine and to search corpora of texts and/or pictures e. g. on the field of archaeology, in the historical linguistics or in the art science,

- To annotate or to comment textual material in literary studies of the classical philology until the letters in modern times,

- To enrich and to link documents with additional materials from different media in history, in the editorial science or in lexicographical studies.

You will find many more examples, when you look at the various projects of the digital humanists worldwide concerning nearly all subjects of the library affine humanities and of the social sciences too. From that background we can observe a development of data driven research on research fields, where you find mainly text oriented research before. What has been read and evaluated as texts until now, will be processed and analyzed as data sets. By this the understanding of knowledge is going to be changed into a data-based empiricism. This pushes the text-based approach of even those disciplines in the background whose performance has so far been rather theoretically or hermeneutically oriented. Therefore the new approaches of the Digital Humanities establish the empiricism which is obtained from data as new, contemporary paradigm of knowledge in the humanities research.

3. Enhanced Publications

Now we come back to the matter of publications, which are affected and influenced by these developments. Because only the given shapes of texts are not enough in order to publish or to represent the findings of data driven research. Against that background we observe the development of so called enhanced publications , which are covering these additional needs. The data, which researchers produce and work within their projects, become part of the research results, so that the data will be part of the published results as well. Examples of

enhanced publication in scholarly communication are available yet. From the many examples some quite characteristic publications are mentioned here:

- Shotton, D.; Portwin, K.; Klyne, G. and Miles, A.: Adventures in semantic publishing: exemplar semantic enhancement of a research article. PLoS Computational Biology 5 (4), (2009): e1000361. http://www.ploscompbiol.org/doi/pcbi.1000361

- Journal of Visualized Experiments: http://www.jove.com/

- Self-Organizing Properties of Mouse Pluripotent Cells Initiate Morphogenesis upon Implantation, by Ivan Bedzhov and Magdalena Zernicka-Goetz: http://www.cell.com/fulltext/S0092-8674%2814%2900075-0#Summary and http://dx.doi.org/10.1016/j.cell.2014.01.023

- VECTORS. Journal of Culture and Technology in a Dynamic Vernacular: http://vectors.usc.edu/journal/index.php?page=Introduction and http://vectors.usc.edu/projects/index.php?project=99

Obviously many different shapes of enhanced publications can be watched in the various disciplines. Even so their basic patterns can be classified as following:

- **Enriched:** Digital publications, which can be classified as enriched publications, include beyond the text or the textual components different data sets from various resources, which are worked out or used by researchers or by research groups. These resources may be animations, audios, digitized pictures or texts, images, measured data, photographies, movies, simulations and even software tools. All these materials or components are aggregated in one document and can be stored as a digital object .

- **Social:** If we discuss digital publications or if we ask questions of understanding to single passages or words directly in the document, we communicate and/or interact with tools, which are known from the WEB 2.0 or form the social networks . From this point of view digital documents or objects become social . This can be made by annotations, chats, comments or any other kinds of likes and dislikes . Social aspects of publications may concern the discussion during the pre-phase of a publication, but the post-phase as well. Mostly we observe annotations and comments after the research results were published. However social facilities might also be used in the peer reviewing process.

- **Processible:** Materials, which based on machine readable data or have been converted or transformed to machine readable data sets, can be exploited or processed for many purposes like data or text mining or searching materials and resources. Therefore enhanced publications have to be structured in order to become processible and searchable. This includes the semantic possibilities of structuring too. Based on RDF-structured data and texts completely new opportunities of combining and linking of data will be opened up even under the aspects of enhanced publication .

- **Linked:** In former times people said, that everything is unpublished, which doesn t hold an ISBN or an ISSN. What the registration of standard book or serial numbers

mattered for printed publications, is the WEB in the digital paradigm: Nothing exists outside the WEB! Digital publication must be accessible and available in the WEB with other words: enhanced publications and as well their singles components must be properly formatted for and delivered to various links and/or platforms.

The classification above is closely connected to the single procedures of the publication process, which will be transformed in a networked process now. The new and interactive shape of the publication process will renew the relationships and roles between the players of the value chain. The traditional players such as publishers, booksellers and libraries are developing either new strategies in the market of scientific information or they will lose their previous position in the value chain. New players such as media and computer centers and commercial IT providers (e. g. software developers, search engines providers) will be added. It is also necessary to design a number of new features within the value chain of publishing, which is not a chain any more, but rather a network for publishing research results.

4. Future Knowledge Management

The focal point will be the new challenges for the libraries in cooperation with the media and the computer centers of the universities. In the years 2011 and 2012, the German Research Society and the German Research Council published their recommendations for the further development of academic infrastructures to improve research conditions towards digital research environments in the current decade until 2020. The general topics of these papers are national licenses, hosting and long term archiving of materials, open access publishing and repositories, digitization of cultural heritage items, research data management, virtual research environments, e-competence and e-literacy. All these measures have an immediate effect on the research activities and the research results and on the publications themselves. They are necessary for the research processes as well as for the publication of the research results. In other words: If they are neglected in the publication of research results, all potentials of the information infrastructures will be wasted.

This leads us to the conclusion that we will implicitly have to integrate the different means of information infrastructures in all phases of research work and finally in the enhanced - publications of the research results. But what are the specific challenges of knowledge management in future? What do we face, if libraries deal with publications, which are aggregations of data and text sets and not only textual materials anymore? It might be surprising: But their future tasks bring libraries back to their roots as archives for the record of human knowledge however comparing with former times it is extremely different now! Considering the change of paradigm from texts towards data we are in the situation to redesign our tasks in depth and to align them to the digital paradigm of data driven research. Therefore we have to care for long term archiving of the many different texts and data types and to record these materials with persistent identifiers and by various metadata forms. Otherwise they won t be searchable in the WEB and will be lost. We have to prepare the materials for further use cases and to provide the necessary requirements to link, to process and to share all of it. We have to cooperate with each other and with the researchers to manage these new challenges, because none of us is able to do this work alone. As a result we have to build up networked collections of data and texts as WEB-based hubs or platforms for all the materials our researchers create and publish.

The following issues are part of the project Future Publications in the Humanities and will find out the demands and needs of knowledge management in depth:

- **Patterns of the research work today**: The representatives of selected disciplines of the humanities will be surveyed on the basis of expert interviews. The questions of these interviews concern in particular the usage of research resources in research projects, the approaches and methods, which are used primarily for dealing with project related data sets, and the inclusion of data and texts in the publication of the research results.

- **Future roles of the information infrastructure:** The representatives of institutions of the information infrastructures (bookseller, libraries, media and computer centers as well as publishers and commercial IT providers) will be consulted in expert interviews too. The interviews should identify the new tasks and roles that will arise from enhanced publications like the preparation, navigation, structuring and visualization of data and contents including enrichments, annotations and options for their re-use. In addition it has to be determined, how the academic support processes of libraries, media and computer centers must be aligned to the new tasks like digitization, research data management, hosting, long-term archiving, open access publishing, repositories, editing and structuring of data etc.

- **New demands for knowledge management:** On the results of the both surveys a future model for knowledge management will be found, which covers the creation, the use and the long-term availability of enhanced publications . From that the organizational and technical interdependencies can be identified. The intended model is based on the following three components which react on each other: (1) preparation and navigation of data and contents as results and/or processes (editing, design and visualization) (2) service and support, which the players of the publication process and/or the institutions of the academic information infrastructures provide, referring to technical formats, uniform identifiers and interoperability (3) business and exploitation models including the legal protection of access and re-use of data and contents.

This means a decisive change in the library service portfolios. The traditional range of tasks which was merely confined to collecting, indexing and providing publications on a long-term basis is significantly extended by added-value-services in order to support the research process on the whole. This is probably the most crucial change that is brought about by the transition from the traditional to the digital libraries. Unlike those services which are in their structure geared to printed contents and texts, data and their handling are gaining centre stage in digital libraries. Data and data stocks even from our familiar materials like digital or printed books and journals - have to be collected, processed, contextualized, structured, linked, hosted, archived, and last but not least, prepared for retrieval and re-use. This development is a great chance and a big challenge for the libraries and they will have to accept them. Their service portfolios will have to cater for research processes and research findings which do not focus on the text paradigm but mainly on data with their particular structure information. At the same time, the generation, distribution and use of data-based knowledge and the necessary infrastructure will not be seen as two different worlds anymore but as interconnected parts.

5. Outlook

If libraries continue the traditional paths, they will further lose awareness, because the knowledge management concerning printed books and licensed journals will not increase their impact. Platforms like Elsevier s SciVerse or Springer Link might be even better than

anything, what libraries will offer one day. What about Google and the Google Apps? Are we ready to compete with Google and to win this battle? From that libraries should better focus on the unique selling point of providing digital data and materials, which are owned by their users or by the research institutions libraries belonging to. We started already with open access publishing and we should carry on this way with a broader scale. That is a big step, but it is our future task we can agree upon: Yes, we can !

6. Bibliography

Borgman, Christine L.: Data, disciplines, and scholarly publishing. in: Learn. Pub 21 (2008), p. 29 38

Büttner, Stephan; Hobohm, Hans-Christof; Müller, Lars (Hg.): Handbuch Forschungsdatenmanagement. Bad Honnef: Bock + Herchen, (2011) p. 35 48 - electronic version: http://opus.kobv.de/fhpotsdam/volltexte/2011/227/pdf/

Current Models of Digital Scholarly Communication by order of the Association of Research Libraries (ARL) - Results of an Investigation Conducted by Ithaka for the Association of Research Libraries, (November 2008), Online-Zugriff: http://www.arl.org/bm~doc/current-models-report.pdf

Degkwitz, Andreas: What will future publications be like? - in: Informationswissenschaft zwischen virtueller Infrastruktur und materiellen Lebenswelten : Proceedings des 13. Internationalen Symposiums für Informationswissenschaft (ISI 2013), Potsdam, 19. bis 22. März 2013 / ed. by Hans-Christoph Hobohm. - Glückstadt: Hülsbusch, 2013, p. 81 - 92 - electronic version: http://opus4.kobv.de/opus4-fhpotsdam/frontdoor/index/index/docId/399

Dudek, Sarah: Die Zukunft der Buchstaben in der alphanumerischen Gesellschaft. Text und Dokument unter digitalen Bedingungen. in: Bibliothek, Forschung und Praxis 36 (2012), p. 189 199

Fournier, Johannes: Zugang, Nachnutzung und Reproduzierbarkeit. Anmerkungen zur künftigen Ausrichtung einer wissenschaftsadäquaten Informationsinfrastruktur. in: Bibliothek Forschung und Praxis 36 (2012), p. 180 188

Füssel, Stephan (hrsg.): Medienkonvergenz Transdisziplinär. - Berlin, de Gruyter (2012), p. 1 6

Giaretta, David; Hudson, Richard L. (Hg.): Riding the wave: How Europe can gain from the rising tide of scientific data. Final report of the High Level Expert Group on Scientific Data. A submission to the European Commission (2010). Unter Mitarbeit von John Wood, Thomas Andersson, Achim Bachem, Christoph Best, Françoise Genova, Diego R. Lopez et al. Online-Zugriff http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf

Gradmann, Stefan; Meister, Jan Christoph: Digital Document and Interpretation. Re-Thinking Text and Scholarship in Electronic Settings. - in: Poiesis & Praxis. International Journal of Ethics of Science and Technology Assessment 5/2 (2008), p. 139 - 153

Groth, Paul; Gibson, Andrew; Velterop, Jan: The anatomy of a nanopublication. in: Information Services and Use 30 (2010), p. 51 56 electronic version http://iospress.metapress.com/content/ftkh21q50t521wm2/

Hogenaar, Arjan; Tjalsma, Heiko; Priddy, Mike: Research in the Humanities and Social Sciences. Chapter E. in: Christian Meier zu Verl und Wolfram Horstmann (Eds.): Studies on

subject-specific requirements for open access infrastructure. 1st edition, Bielefeld: Univ.-Bibliothek, (2011) p. 165 214

Key Perspectives Ltd.: Data Dimensions: Disciplinary Differences in Research Data Sharing, Reuse and Long term Viability. DCC SCARP Synthesis Report. Digital Curation Centre (DCC) (2010). - http://www.dcc.ac.uk/sites/default/files/SCARP%20SYNTHESIS_FINAL.pdf.

Lynch, Clifford: The Shape of the Scientific Article in The Developing Cyberinfrastructure. in: CT-Watch Quarterly 3 (2007). Online-Zugriff http://www.ctwatch.org/quarterly/articles/2007/08/the-shape-of-the-scientific-article-in-the-developing-cyberinfrastructure/

Lyon, Liz; Rusbridge, Chris; Neilson, Colin; Whyte, Angus: DCC SCARP: Disciplinary Approaches to Sharing, Curation, Reuse and Preservation. JISC Final Report. JISC; Digital Curation Centre (DCC) (2010). Electronic version http://www.dcc.ac.uk/sites/default/files/documents/scarp/SCARP-FinalReport-Final-SENT.pdf

Mittler, Elmar: Wissenschaftliche Forschung und Publikation im Netz. - in: Medienkonvergenz - Transdisziplinär. ed. by Stephan Füssel. - Berlin, de Gruyter, (2012), p. 32 80

Neuroth, Heike; Strathmann, Stefan; Oßwald, Achim; Scheffel, Regine; Klump, Jens; Ludwig, Jens (Ed.): Langzeitarchivierung von Forschungsdaten. Eine Bestandsaufnahme. Boizenburg, Göttingen: Hülsbusch; Univ.-Verl. Göttingen (2012). Electronic version http://www.nestor.sub.uni-goettingen.de/bestandsaufnahme/nestor_lza_forschungsdaten_bestandsaufnahme.pdf

Purdy, James P.: The Changing Space of Research: Web 2.0 and the Integration of Research and Writing Environments. in: Computers and Composition 27, (2010) p. 48 58

Seadle, Michael: Archiving in the Networked World. Interoperability. - in: Library Hi Tech 28, (2010)

Verhaar, Peter: Report on Object Models and Functionalities. - DRIVER, Digital Repository Infrastructure Vision for European Research II. Leiden (2008), Online-Zugriff: http://www.driver-repository.eu/component/option,com_jdownloads/Itemid,83/task,summary/cid,54/catid,8/

Woutersen-Windhouwer, Saskia; Brandsma, Renze: Report on Enhanced Publications state-of-the-art. - DRIVER, Digital Repository Infrastructure Vision for European Re-search II. Amsterdam, (2009) Online-Zugriff: http://www.driver-repository.eu/component/option,com_jdownloads/Itemid,83/task,summary/cid,53/catid,8/